

# The effect of perceptual learning on the sensitivity and stability of

## double fusion in Panum's Limiting Case

Yuyu Shi<sup>1</sup>, Yanyan Gao<sup>1</sup>, Jing Ren<sup>1</sup>, Ashley Chung-Fat-Yim<sup>2</sup>, Huayun Li<sup>1,\*</sup> & Xize Jia<sup>1</sup>

<sup>1</sup> College of Teacher Education, Zhejiang Normal University, Jinhua, China

<sup>2</sup> Department of Psychology, York University, Toronto, Canada

### Author Contributions

All of the authors contributed to the design of the study and preparation of the manuscript. Yuyu Shi proposed the experiment design. Yanyan Gao and JingRen acquired and analyzed the data. Ashley, Huayun Li and Xize Jia contributions to the modification of the article.

\*Corresponding author: Huayun Li (E-mail: lihuayun99@163.com)

Running title: The effect of training on double fusion

### Abstract

Panum's limiting case is a typical phenomenon of monocular occlusion in binocular vision. It occurs naturally when one object is occluded by another object for one eye, but in the other eye the two objects are located in different directions. Only recently has it been found that in addition to horizontal disparity, the vertical disparity gradient and cue conflict are two important determinants for double fusion. Therefore, the current study aims to determine the relationship between the two determinants and perceptual learning in Panum's limiting case. Twenty-six observers were trained for five days. Meanwhile, the RTs and duration of double fusion was measured when the participants viewed several versions of Panum's configuration. In these stimuli, vertical disparity gradient was varied from 0.1 to 0.6 and cue conflict was manipulated from low to high. The results revealed that for each level of these factors, the RTs of double fusion decreased and the duration of double fusion increased with each training session. Moreover, there were significant differences among different levels of vertical disparity gradient and cue conflict. Lastly, there was also a significant interaction effect between the two determinants in Panum's limiting case. These results suggest that there is perceptual learning for each level of the two determinants in monocular occlusion and these factors jointly affect the sensitivity and stability of double fusion.

**Keywords:** monocular occlusions, Panum's limiting case, vertical disparity gradient, cue conflict, double fusion, training.

## 1 Introduction

When the world is viewed binocularly, the visual system receives two slightly different images of the scene projected to each eye. Stereopsis is the process responsible for reconstructing three-dimensional (3D) depth perception from these two different two-dimensional (2D) images (Wheatstone, 1838). In natural scenes, an object may appear to occlude another object in one eye, resulting in only one perceived object in this eye, but in the other eye the two objects will be visible at two different distances. Under this condition, the monocular area in the two eyes arise and it is referred to as "monocular occlusion" (Kaye, 1978; Tsirlin et al., 2014). Since the advent of the stereoscope by Wheatstone in 1838, the majority of stereoscopic research has focused on understanding how binocular disparity is used in depth perception (Harris & Wilcox, 2009; Howard & Rogers, 2012). But there was little interest in the role of the occluded regions, which were treated as noise when reconstructing 3D images. Only in the past two decades has there been a growing interest in demonstrating that the visual system could estimate depth in the occlusion (Anderson & Julesz, 1995; Gillam & Borsting, 1988; Nakayama & Shimojo, 1990).

One of the simplest phenomena to include monocular occlusion in binocular depth perception is Panum's limiting case (Gillam et al., 1995; Panum, 1858). It arises when one eye is presented with a single feature, while the other eye is presented with two horizontally offset features. Such a configuration is consistent with a pair of real objects located at different distances (see Fig. 1): the closer object occludes the more distant one for the left eye but not for the right eye. The depth mechanism of Panum's limiting case is important in

establish the theory of stereopsis, however, it has long been the subject of controversy in binocular vision (Ono et al., 1992).

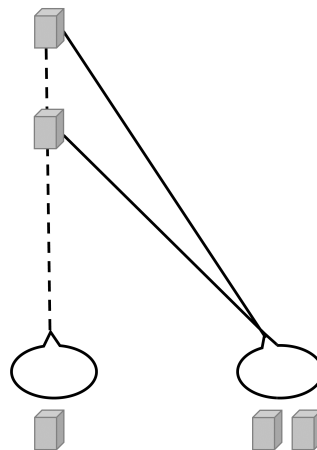


Fig .1. An example of Panum's limiting case

The debate in Panum's limiting case stems from whether the perceived depth is derived from double fusion or single fusion. On one hand, the results supported for double fusion showed that when the disparity is small, depth was perceived by matching the single feature presented to one eye with the two features presented to the other eye (Gillam et al., 1995; Gillam et al., 2003; Kumar, 1996). On the other hand, the advocates of single fusion proposed that only one feature was derived from binocular matching and the other one originated from other cues (Frisby, 2001; Ono et al., 1992; Shimono et al., 1999; Wang et al., 2001). Only recently has it been discovered that in addition to disparity, the vertical disparity gradient and cue conflict are two important factors to the final depth perception of Panum's configuration (Li et al., 2012). The vertical disparity gradient refers to the ratio of the difference between the minimum and maximum disparity point pairs of the two eyes in the configuration to the binocular septum (Bülthoff et al., 1991; Burt & Julesz, 1980). Cue conflict (see Figure 2) refers to the conflict between various cues that provide information sources for depth perception. In this study, cue conflict refers to the inconsistency between the two-dimensional shape provided by monocular cues and the three-dimensional shape provided by binocular parallax (Allison & Howard, 2000). Moreover, recent evidence by Li and colleagues has shown that in perception, double fusion always occurred in the process of stereopsis (H. Li et al., 2011). Hence, the effect of vertical disparity gradient and cue conflict on double fusion was examined in this study.

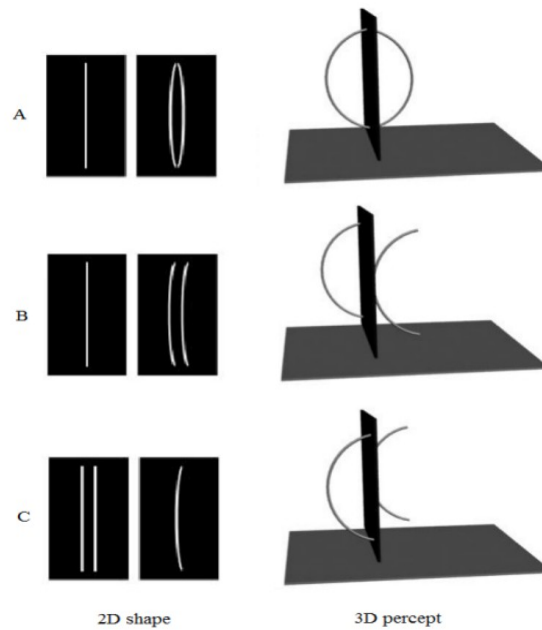


Fig .2. Schematic illustrating the cue conflict in Gilliam series. On the left is the two-dimensional image provided by the monocular cues, and on the right is the 3D image after perception. From top to bottom, the matching degree between 2D and 3D images decreases, and the cue conflict increases.

It is well known that perceptual learning can generate performance improvements due to repeated exposure and practice to sensory tasks. Such training-induced plasticity has been established in a variety of visual tasks, ranging from the discrimination of simple attributes to more complex sensory tasks (Hussain et al., 2009; Tsodyks & Gilbert, 2004). The simple tasks involve orientation identification, direction discrimination, or motion detection (Ball & Sekuler, 1987; Huang et al., 2007; J. et al., 2015; Shiu & Pashler, 1992). Examples of complex tasks involve face identification, video game playing, or contour judgment (Gold et al., 1999; R. W. Li et al., 2011; Mckendrick & Battista, 2013). Many studies have demonstrated that there are both rapid and slow learning in the time course of perceptual learning, but it is not sure about the time course in Panum's limiting learning. And it remains unclear what the relationship is between perceptual learning and the determinants to stereopsis in Panum's limiting case.

Previous studies on the Panum's limiting case mainly focused on the types or forms of stimulus, but paid little attention to the following aspects, and there was also a lack of further research on the perception learning of stereoscopic vision.

To be specific, first, previous studies rarely considered the effect of stereoscopic experience on depth perception outcomes. Some researchers used both experienced and non-experienced subjects in an experiment without making a distinction between their results (Bingushi & Yukumatsu, 2010; Wang et al., 2001). While experienced subjects were more likely to perceive double fusion, non-experienced subjects were more likely to perceive single fusion. So the results of a true double fusion may be influenced. Second, when the stimulus was presented for a short time, the perceptual results of the subjects were not stable. Even when presented with the same stimulus, subjects may perceive distinct results at

different times. Moreover, the perceptual result of double fusion takes some time (Spang et al., 2012). Therefore, if the presentation time is short, the subjects may be more inclined to perceive the single fusion. Third, in previous studies, some studies have looked at the effect of factors on perceptual outcomes. Subjects were asked to report whether they perceived double fusion during the experiment. The number of double fusions perceived by all subjects to this stimulus was used as the dependent variable (Wang et al., 2001). However, as the frequency is discrete data and not normally distributed, the parameter test cannot be carried out and the interaction between these influencing factors can't be analyzed.

This study aims to investigate the relationship between the two determinants (vertical disparity gradient, cue conflict) and perceptual learning in monocular occlusion. In contrast to previous studies with similar themes, the current study has more strength. In this study, only non-empirical subjects were selected and the stimulus response time was set as 10s to provide the subjects with a longer stimulus response time. And the subjects could press the button multiple times to report their perceptual results at any time. In addition, reaction time and duration were used as dependent variables to further consider whether there was a rapid learning stage in Panum's limiting time. The RTs and duration of double fusion were measured and the observers were trained for five days. Vertical disparity gradient varied between 0.1, 0.3, and 0.6. Cue conflict was manipulated from low, medium, to high. Based on the existing literature, we hypothesized that there would be an improvement in performance for each level of the two determinants after training. The two determinants could affect the sensitivity and stability of double fusion. Additionally, we hypothesized there were significant differences among different levels of these factors.

## 2 Method

### 2.1 Observers

A total of 26 undergraduate and graduate students were selected through open recruitment, which aged 22 to 29 years old. There are 14 males and 12 females. They didn't have any prior experience of stereopsis. All observers with written informed consent had normal or corrected-to-normal visual acuity and good stereoscopic acuity as measured by the Randot<sup>TM</sup> stereoacuity test (at least 20 s of arc). The study was approved by the research ethics board at Southeast University and followed the tenets of the Declaration of Helsinki.

### 2.2 Apparatus

Scripts for stimulus presentation were executed on a G5 Macintosh computer using Matlab version 2.54 of Psychtoolbox extensions (Brainard, 1997; Pelli, 1997). Stimuli were presented on a pair of CRT monitors (ViewSonic G225f) arranged in a mirror stereoscope at a viewing distance of 60 cm. The resolution of the monitors was 1280\*960 pixels with a refresh rate of 75 Hz. Participants were seated in a dark room with their heads stabilized by a chin rest to minimize head movement.

### 2.3 Stimuli

The stimuli all included a single feature presented to one eye and two features presented to the other eye. Vertical disparity gradient varied between 0.1, 0.3, and 0.6. Cue conflict was manipulated from low, medium to high(see Fig. 3). Three kinds of stereoscopic stimuli were included in each experimental condition to depict Panum's limiting case (i.e., series of Gilliam, Frisby, and Wang). A total of 54 kinds of stimuli were constructed.

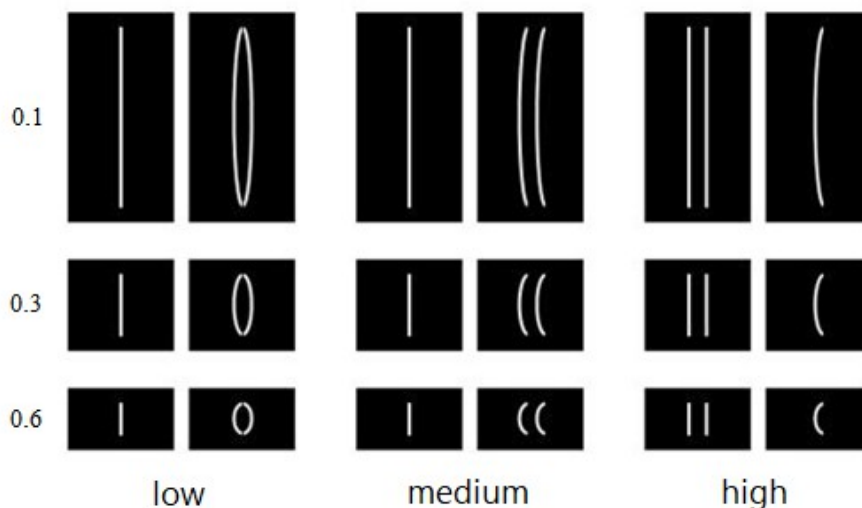


Fig. 3. As shown above is the nine experimental conditions in the Gilliam arc configuration. In the vertical direction, the vertical disparity gradient ranges from 0.1, 0.3 and 0.6. In the horizontal direction, there are three levels of cue conflicts from left to right, low, medium and high.

### 2.4 Design and Procedure

The current study consisted of six sessions: a preliminary session and five training sessions. The instructions for the preliminary and training sessions were identical. In the preliminary session and training sessions, participants were instructed to distinguish between “single” or “double” fusion for each stimulus. Participants were told that there was no correct answer and they only needed to respond through pressing the button on the keyboard whether they see “single” or “double” fusion. Participants were told that double fusion refers to two features both having 3D depth after fusion, while single fusion refers to only one of the two features having 3D depth after fusion. Once participants understood the criteria for single and double fusion, they completed 9 practice trials as the preliminary session. Afterwards, participants completed five training sessions (day1, day2, day3, day4 and day5) that were scheduled five days in a row for about twenty minutes every day. Each training session consisted of the 54 kinds of stimuli.

On each trial (see Fig 4), two rectangles appeared for an undefined amount of time. Their size was the same as the rectangles used for the stimuli. Once participants fused the two rectangles, they were instructed to press the ‘Space’ key to begin the trial. The stimulus (the target features inside the rectangle) then appeared on the screen for 10 seconds. When participants perceived double fusion, they were instructed to press ‘S’ as quickly and

accurately as possible. If participants perceived the perception change to single fusion, they were instructed to press 'F'. During the 10 seconds period, they could press the two keys as many times as they want if their perception changed. If they did not perceive double fusion after the stimulus appeared, they were told not to press any key until 10 seconds had elapsed. No matter how many times the observers pressed the key, the stimulus remained on the screen for 10 seconds. Once the stimulus disappeared, the two rectangles re-appeared and the participant began a new trial.

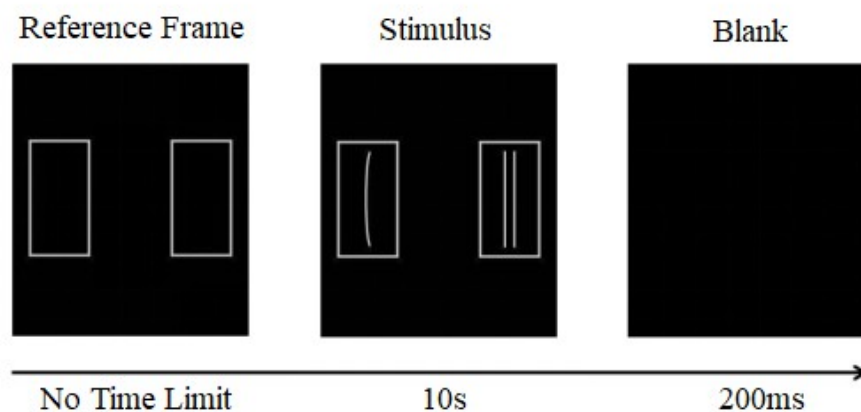


Fig. 4. A picture for the experimental process on each trial

## 2.5 Data analysis

Double fusion for all the observers were calculated separately. If the participant did not press any key, the RT was recorded as 10 seconds and duration as 0 seconds for the analysis. If the perception changed during the 10 seconds, the program recorded the first time double fusion was perceived for RT and calculated the total time of double fusion was perceived for duration.

Statistical analyses were carried out using Statistica 7.0 (StatSoft Inc). Two separate three-way ( $3 \times 3 \times 5$ ) repeated-measures ANOVAs were performed for mean RTs and duration of double fusion. The three factors were vertical disparity gradient (0.1, 0.3, and 0.6), cue conflict (low, medium, and high) and training session (day1, day2, day3, day4, and day5), respectively. A Greenhouse–Geisser correction was applied when sphericity was violated. Post-hoc t-tests (Bonferroni corrected for multiple comparisons) were performed when necessary.

## 3 Results

### 3.1 The results of RTs for double fusion

There was a main effect of vertical disparity gradient,  $F(2, 50) = 46.85$ ,  $p < .001$ ,  $\eta_p^2 = .65$ . Post-hoc paired t-tests revealed that the RT for the disparity gradient of 0.1 ( $M = 3.24$ ,  $SE = .31$ ) was significantly shorter than 0.3 ( $M = 4.12$ ,  $SE = .35$ ) and 0.6 ( $M = 5.31$ ,  $SE = .42$ ), both  $ps < .001$ . The RT for disparity gradient 0.3 was significantly shorter than 0.6,  $p < .001$ .

There was also a main effect of cue conflict,  $F(2, 50) = 41.62$ ,  $p < .001$ ,  $\eta_p^2 = .63$ . Post-hoc paired t-tests revealed that the RT of high cue conflict ( $M = 6.12$ ,  $SE = .45$ ) was significantly longer than medium ( $M = 3.35$ ,  $SE = .37$ ) and low ( $M = 3.21$ ,  $SE = .38$ ),  $ps < .001$ . Moreover, the main effect of training was significant,  $F(4, 100) = 11.45$ ,  $p < .001$ ,  $\eta_p^2 = .31$ . Post-hoc paired t-tests revealed that the RT of day1 ( $M = 4.92$ ,  $SE = .32$ ) was significant longer than day3 ( $M = 4.07$ ,  $SE = .41$ ), day4 ( $M = 3.87$ ,  $SE = .38$ ) and day5 ( $M = 3.64$ ,  $SE = .38$ ), all  $ps < .004$ . The RT of day2 ( $M = 4.63$ ,  $SE = .35$ ) was significantly longer than day3, day4 and day5,  $ps < .01$ .

For the interaction effect, the results revealed that there was a significant interaction between vertical disparity gradient and cue conflict,  $F(4, 100) = 12.88$ ,  $p < .001$ ,  $\eta_p^2 = .34$ . (see Fig 5) For three levels of vertical disparity gradient, the RTs of high cue conflict were all significantly longer than medium and low cue conflict, all  $ps < .001$ . Note that, the 0.1 and 0.3 vertical disparity gradient conditions yielded a marginal difference between low and medium cue conflict,  $ps < .05$ . The RTs of medium conflict were longer than that of low conflict. While for 0.6 vertical disparity gradient, the RTs of medium conflict were all significantly shorter than that of low conflict,  $ps < .02$ .

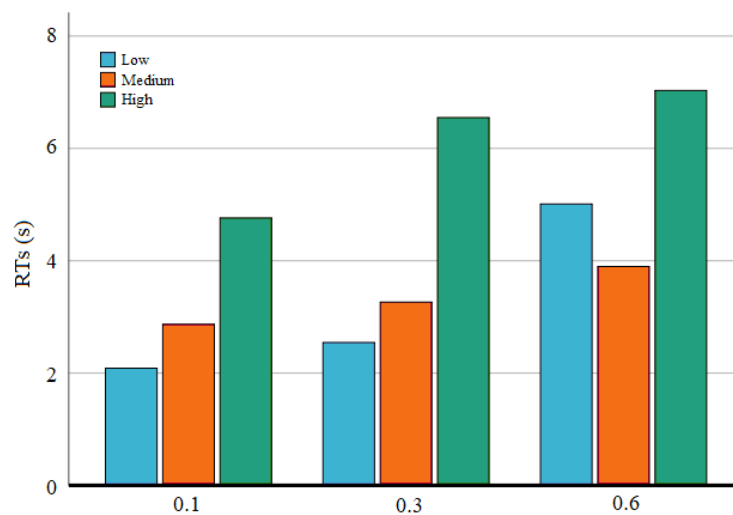


Fig. 5. The RT results for the relationship between cue conflict and the vertical gradient disparity.

The RT results for the two factors (vertical disparity gradient and cue conflict) by training session are presented in Figures 6 and 7, respectively. Although the interaction effects between the two factors and training session were not significant, the RTs of the two factors in any two successive levels all decreased from day1 to day5. With each training session, the RTs for double fusion decreased. Paired sample t-tests confirmed that the RTs of every level of the two factors were all significantly different between day1 and day5, all  $ps < .001$ .



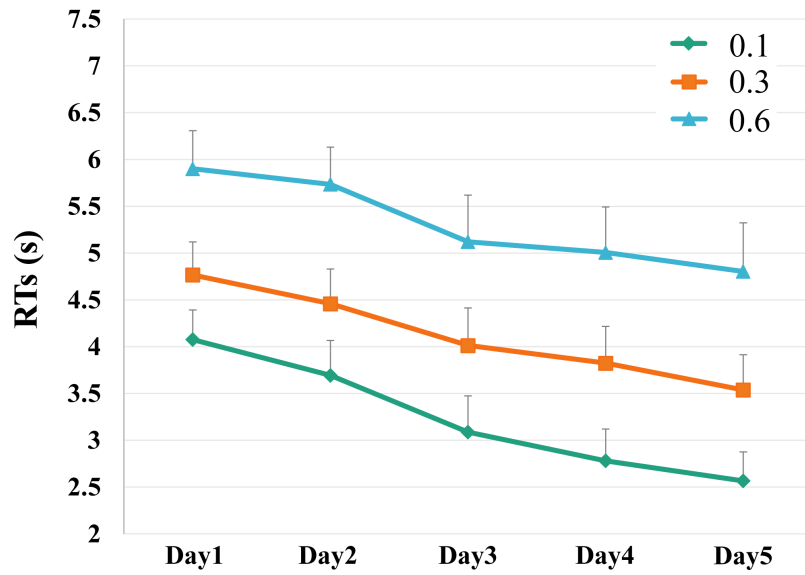


Fig. 6. The RTs results for the relationship between vertical gradient disparity and training. The X-axis represents the five training sessions. The green, orange and blue line represent RTs of vertical gradient disparity (0.1, 0.3, 0.6) respectively. Error bars show 95% confidence interval.

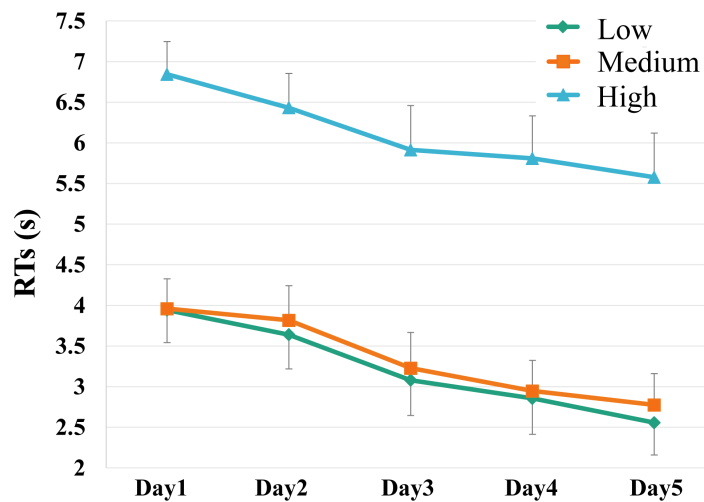


Fig. 7. The RTs results shows the perceptual learning of cue conflict. The X-axis represents the five training sessions. The green, orange and blue line represent RTs of cue conflict (low, medium, high) respectively. Error bars show 95% confidence interval.

### 3.2 The results of duration for double fusion

The three-way ANOVAs on duration of double fusion was performed. There was a main effect of vertical disparity gradient,  $F(2, 50) = 49.52$ ,  $p < .001$ ,  $\eta_p^2 = .67$ . Post-hoc paired t-tests revealed that the duration between each level of disparity gradient were all significant,  $ps < .001$ . Specifically, the duration of 0.1 disparity gradient ( $M = 6.69$ ,  $SE = .31$ ) was significantly longer than 0.3 ( $M = 5.78$ ,  $SE = .35$ ) and 0.6 ( $M = 4.53$ ,  $SE = .42$ ). The duration

of 0.3 disparity gradient was significantly longer than 0.6 disparity gradient,  $p < .001$ . There was also a significant main effect of cue conflict,  $F(2, 50) = 45.01$ ,  $p < .001$ ,  $\eta_p^2 = .64$ . Post-hoc paired t-tests revealed that the duration of high cue conflict ( $M = 3.72$ ,  $SE = .44$ ) was significantly shorter than medium ( $M = 6.56$ ,  $SE = .36$ ) and low cue conflict ( $M = 6.73$ ,  $SE = .38$ ),  $ps < .001$ . Moreover, the main effect of training was significant,  $F(4, 100) = 10.43$ ,  $p < .001$ ,  $\eta_p^2 = .29$ . Post-hoc paired t-tests revealed that duration of day1 ( $M = 4.98$ ,  $SE = .31$ ) was significant shorter than day3 ( $M = 5.84$ ,  $SE = .40$ ), day4 ( $M = 6.03$ ,  $SE = .38$ ) and day5 ( $M = 6.22$ ,  $SE = .38$ ), all  $ps < .01$ . The RT of day2 ( $M = 5.28$ ,  $SE = .35$ ) was significantly shorter than day 3, day4 and day5,  $ps < .05$ .

For the interaction effect, the results revealed that there was a significant interaction between vertical gradient and cue conflict,  $F(4, 100) = 12.58$ ,  $p < .001$ ,  $\eta_p^2 = .34$ . No other interaction effect was significant,  $ps > .05$ .

The duration results of relationship between the two factors (vertical disparity gradient and cue conflict) and training session are shown in Figures 8 and 9 respectively. Although the interaction effects between the two factors and training session were not significant, the duration of the two factors in two successive levels increased from day1 to day5. As training session progressed, the duration of double fusion increased. Paired t-tests confirmed that for the duration of every level of the two factors, there were all significant differences between day1 and day5, all  $ps < .001$ . The duration of day1 was all significantly shorter than that of day5 at every level of the two factors.

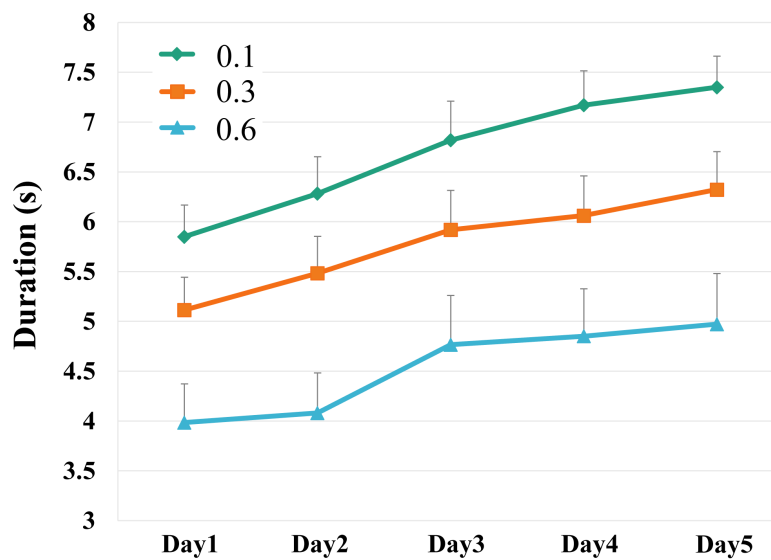


Fig. 8. The Duration time results for the relationship between vertical gradient disparity and training. The X-axis represents the five training sessions. The green, orange and blue line represent RTs of vertical gradient disparity (0.1, 0.3, 0.6) respectively. Error bars show 95% confidence interval.

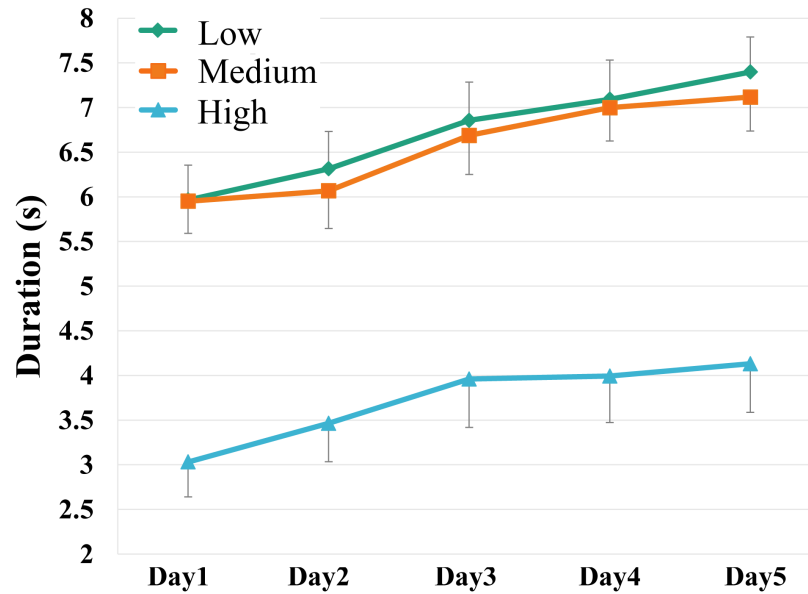


Fig. 9. The Duration time results shows the perceptual learning of cue conflict. The X-axis represents the five training sessions. The green, orange and blue line represent RTs of cue conflict (low, medium, high) respectively. Error bars show 95% confidence interval.

#### 4 Discussion

In this study, we examined the relationship between the two determinants of stereopsis (vertical disparity gradient and cue conflict) and perceptual learning. The results showed that for each level of the two factors, the RTs and duration of double fusion for day1 were all significantly different from day5, providing evidence that perceptual learning occurred in the two determinants. Moreover, the results showed that for RTs and duration of double fusion, the main effects of vertical gradient and cue conflict were significant. Additionally, the interaction effect between vertical gradient and cue conflict on double fusion also reached significance. These results suggest that vertical disparity gradient and cue conflict are two important factors to depth perception in Panum's limiting case, and the two determinants jointly affect the sensitivity and stability of double fusion in stereopsis.

##### 4.1 The relationship between vertical disparity gradient and perceptual learning

Consistent with previous studies that have demonstrated the role of vertical disparity gradient on stereoscopic fusion (Bülthoff et al., 1991; Burt & Julesz, 1980; Mckee & Verghese, 2002; Yu et al., 2017), the present study corroborated and extended previous results by showing that the smaller the vertical disparity gradient, the shorter the RTs for double fusion but the longer the duration of double fusion. Additionally, compared to the study of Li et al. (Li et al., 2012), which used the frequency of reported double fusion to investigate this determinant, this study developed a more innovative indicator to examine the effect of vertical gradient on depth perception. Specifically, the results for the RTs and duration of double fusion in this study performed the three-way repeated measures ANOVAs, which could evaluate the interaction effect among these factors. However, the results for the

frequency of double fusion can only be analyzed by the non-parametric test, e.g., Cochran's Q test and McNemar's test, which is difficult to evaluate the interaction effect among more than two factors (Nelder, 1965; Somes & Bhapkar, 1977)

It is widely accepted that perceptual learning can generate visual performance improvements in a variety of visual tasks (Ball & Sekuler, 1987; Huang et al., 2007; O'Toole & Kersten, 1992; Zhou et al., 2006). In our study, the perceptual learning for the sensitivity and stability of double fusion was examined by using stimuli that depicted Panum's limiting case. The results revealed that for all levels of vertical gradient, the RTs of double fusion for day1 were all significantly longer than that of day5, but shorter for duration of double fusion. As training session changed from day1 to day5, the RTs of double fusion for each level of this factor decreased gradually, whereas duration of double fusion increased gradually. These results suggest that there perceptual learning occurred for each level of vertical gradient.

#### *4.2 The relationship between cue conflict and perceptual learning*

Consistent with our hypothesis for cue conflict, there was a significant main effect among the three levels. Further analyses revealed that the RTs and duration of double fusion, the high conflict was significantly different from medium and low conflict, suggesting that the manipulation of cue conflict was effective. Previous research demonstrated that cue conflict could affect depth perception (Pizlo et al., 2005). Moreover, the results also show that for each level of cue conflict, the RTs for double fusion gradually decreased, while the duration gradually increased when the training session changed from day1 to day5, suggesting that perceptual learning did occur in each level of cue conflict.

Based on the results from the relationship between cue conflict and perceptual learning, it is worth mentioning that the result of high conflict is significantly different from medium and low conflict, but medium and low cue conflict are quite similar to each other. Although further analyses reveal a difference between medium and low cue conflict in the two-way interaction effect, the main effect between medium and low cue conflict was not significant. These results are interesting and deserve to be further examined in future study. In fact, Li and colleagues (2012) have realized that the conflict degree of medium is just a little higher than that of low conflict. The results of the present study are consistent with the previous study, jointly illustrating that the manipulation of the three levels for cue conflict are not equidistant. The conflict degree of low is quite different from that of medium and low conflict, but the medium cue conflict is quite similar with that of low conflict.

#### *4.3 The effects and time course of perceptual learning in Panum's limiting case.*

It was found that in all levels of vertical disparity gradient and cue conflict, the perceived response to double fusion decreased gradually with the increase of training days, while the duration increased gradually. This indicates that there are significant perceptual learning effects for the level of vertical disparity gradient and cue conflict. In the Panum's limiting case, perceptual learning is mainly the learning of factors. Learning in the vertical disparity gradient and cue conflict of the stimulus configuration reflects the specificity of

perceptual learning. That is, the training effect achieved on a certain stimulus after changing the orientation or position of the stimulus, it needs to be practiced again to achieve the previous training effect. (Poggio et al., 1992) However, transfer was shown at different levels of vertical disparity gradient and cue conflict, as well as the later training day, in which the previous training effect was retained or the learning speed was accelerated when learning the next task (Ahissar & Hochstein, 1997; Liu & Weinshall, 2000).

Previous studies have suggested that there are two stages of fast learning and slow learning in the time course of perceptual learning, but there is no obvious fast learning stage in this study. There may be two reasons. On the one hand, the task of this study is relatively difficult, so the rapid learning stage is not significant (Avi et al., 1997). On the other, previous studies on the minimum training amount in perceptual learning have put forward the concept of critical amount, believing that only when the critical amount is satisfied in perceptual learning can learning effects be produced (Wright & Sabin, 2007). However, the critical mass in different tasks or stimuli is different, and in this study, there may be no significant rapid learning stage due to the small amount of training.

#### *4.4 The effects of vertical disparity gradient and cue conflict on stereopsis*

The results of this study show that the main effects of vertical disparity gradient and cue conflict are significant, indicating that the two factors are crucial to the influence of double fusion, this is consistent with previous research. Different from previous studies, the response time and duration were taken as dependent variables in this study, and further analysis showed that the interaction between vertical disparity gradient and cue conflict was significant. Specifically, when the vertical disparity gradient was 0.1 and 0.3, the higher the degree of cue conflict, the lower the average response time of double fusion, and the shorter the duration of double fusion. However, when the vertical disparity gradient was 0.6, there was no consistent regularity in the effects of the degree of cue conflict on the average response time and duration of double fusion. It may be that under the condition of 0.6 vertical parallax gradient, the subjects could hardly perceive the double fusion, so the experimental results under this condition could not accurately measure the impact of clue conflict on the double fusion.

#### *4.5 Limitation and future direction*

Although some interesting conclusions have been drawn, a limitation worth to mention involved the response coding. As mentioned earlier, if the participant could not perceive double fusion during stimulus presentation, the RTs were recorded as 10 seconds and duration as 0 seconds. Although this would not affect the trends of each factor, it may lead to the difference being underestimated among mean value of different experimental conditions. For example, if the results of RTs and duration for just several observers are very different from that of most observers, the mean results of RTs and duration for all observers will be different from that of most observers. It may lead to mean RTs becoming longer or mean duration

becoming shorter. Future studies should consider setting a longer duration than 10 seconds. The frequency of recorded for 10 seconds can be found in the Supplementary Materials.

Future direction is to examine the neural correlated of the two determinants on perceptual learning. Previous studies have revealed that when the vertical gradient become lower, the amplitude of ERP components (N170) would become higher (Li, Jia, Chung-Fat-Yim, Jin, & Yu, 2017) and the omega complexity, which could assess the degree of global synchronization between spatially distributed brain areas, would become lower (Li, Jia, & Yu, 2017). The results of this study have shown that when the vertical disparity gradient become lower, the results for double fusion become more sensitive and more stable. It is rational to speculate that the sensitivity and stability of double fusion has certain relationship with ERP components and the omega complexity.

## **5 Conclusion**

This study provides evidence to demonstrate the relationship between perceptual learning and the two determinants in stereopsis, i.e., vertical disparity gradient and cue conflict. The results revealed that for each level of these factors, the RTs of double fusion decreased and the duration of double fusion increased with each training session. Moreover, vertical disparity gradient and cue conflict had an effect on RTs and duration of double fusion, and the interaction between vertical gradient and cue conflict was also significant, suggesting that there is a perceptual learning for each level of the two determinants in stereopsis and these factors jointly affect the perceptual learning. In consequence/Hence, the perceptual learning could improve the sensitivity and stability of double fusion in monocular occlusion.

## **ACKNOWLEDGEMENTS**

This work was supported by the Natural Science Foundation of China under Grant Nos. 61673113, 61273224 and 61074126, and by China Scholarship Council (CSC) under Grant No. 201406090145.

## References

- Ahissar, M., & Hochstein, S. (1997). Task difficulty and the specificity of perceptual learning. *Nature*, 387(6631), 401-406.
- Allison, R. S., & Howard, I. P. (2000). Temporal dependencies in resolving monocular and binocular cue conflict in slant perception. *Vision research*, 40(14).
- Anderson, B. L., & Julesz, B. (1995). A theoretical analysis of illusory contour formation in stereopsis. *Psychological Review*, 102(4), 705-743.
- Avi, Karni, and, Giuseppe, & Bertini. (1997). Learning perceptual skills: behavioral probes into adult cortical plasticity. *Current Opinion in Neurobiology*.
- Ball, K., & Sekuler, R. (1987). Direction-specific improvement in motion discrimination. *Vision Research*, 27(6), 953-965.
- Bingushi, K., & Yukumatsu, S. (2010). Disappearance of a monocular image in Panum's limiting case. *Japanese Psychological Research*, 47(3), 223-229.
- Brainard, D. (1997). The psychophysics toolbox. *Spatial Vision*, 10.
- Bülthoff, H., Fahle, M., & Wegmann, M. (1991). Perceived depth scales with disparity gradient. *Perception*, 20(2), 145-153.
- Burt, P., & Julesz, B. (1980). A disparity gradient limit for binocular fusion. *Science*, 208(4444), 615-617.
- Frisby, J. P. (2001). Limited understanding of Panum's limiting case. *Perception*, 30(9), 1151-1152.
- Gillam, B., Blackburn, S., & Cook, M. (1995). Panum's limiting case: double fusion, convergence error, or 'da Vinci stereopsis'. *Perception*, 24(3), 333.
- Gillam, B., & Borsting, E. (1988). The role of monocular regions in stereoscopic displays. *Perception*, 17(5), 603.
- Gillam, B., Cook, M., & Blackburn, S. (2003). Monocular discs in the occlusion zones of binocular surfaces do not have quantitative depth--a comparison with Panum's limiting case. *Perception*, 32(8), 1009-1019.
- Gold, J., Bennett, P. J., & Sekuler, A. B. (1999). Signal but not noise changes with perceptual learning. *Nature*.
- Harris, J.M., & Wilcox, L. (2009). The role of monocularly visible regions in depth and surface perception. *Vision Research*, 49, 2666-2685.
- Howard, I. P., & Rogers, B. J. (2012). *Perceiving in Depth, Volume 2: Stereoscopic Vision*.
- Huang, C., Tao, L., Zhou, Y., & Lu, Z. L. (2007). Treated amblyopes remain deficient in spatial vision: A contrast sensitivity and external noise study. *Vision Research*, 47(1), 22-34.
- Hussain, Z., Sekuler, A. B., & Bennett, P. J. (2009). Perceptual learning modifies inversion effects for faces and textures. *Vision Research*, 49(18), 2273-2284.
- J., Zhou, F., Yan, Z.-L., Lu, Y., Zhou, J., & Xi. (2015). Broad bandwidth of perceptual learning in second-order contrast modulation detection. *Journal of Vision*.
- Kaye, M. (1978). Stereopsis without binocular correlation. *Vision Research*, 18(8), 1013-1022.
- Kumar, T. (1996). Multiple matching of features in simple stereograms. *Vision Research*, 36(5), 675-698.



- Li, H., Xie, Y., Li, S., Xie, S., Gao, C., & Yang, Z. (2012). The effects of vertical gradient of disparity and combination mode of features on the occurrence of double fusion in Panum's limiting case. *Perception*, 41(11), 1299-1314.
- Li, H., Xie, Y., Li, Y., & Li, S. (2011). Slant Effects in Fold-Line Wheatstone-Panum Limiting Case - A Solid Evidence for Double Matching.
- Li, H., & Yu, D. (2018). Perceptual Learning in Monocular Occlusions: Evidence from Panum's Limiting Case. the 2018 7th International Conference,
- Li, R. W., Ngo, C., Nguyen, J., & Levi, D. M. (2011). Video-game play induces plasticity in the visual system of adults with amblyopia. *PLoS Biol*, 9(8), e1001135.
- Liu, Z., & Weinshall, D. (2000). Mechanisms of generalization in perceptual learning. *Vision Research*, 40(1), 97-109.
- Mckee, S. P., & Verghese, P. (2002). Stereo transparency and the disparity gradient limit. *Vision Research*, 42(16), 1963-1977.
- Mckendrick, A. M., & Battista, J. (2013). Perceptual learning of contour integration is not compromised in the elderly. *J Vis*, 13(1), 5.
- Nakayama, K., & Shimojo, S. (1990). Da vinci stereopsis: Depth and subjective occluding contours from unpaired image points. *Vision Research*, 30(11), 1811-1825.
- Nelder, J. A. (1965). A note on Cochran's Q test. *Biometrics*, 21(4), 1008-1010.
- O'Toole, A. J., & Kersten, D. J. (1992). Learning to see random-dot stereograms. *Perception*, 21(2), 227-243.
- Ono, H., Shimono, K., & Shibuta, K. (1992). Occlusion as a depth cue in the Wheatstone-Panum limiting case. *Attention Perception & Psychophysics*, 51(1), 3-13.
- Panum, P. L. (1858). *Physiologische Untersuchungen über das Sehen mit zwei Augen*. Schwesssche Buchhandlung.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10(4), 437-442.
- Pizlo, Z., Y., Francis, & G. (2005). A new look at binocular stereopsis. *VISION RESEARCH -OXFORD-*.
- Poggio, T., Fahle, M., & Edelman, S. (1992). Fast perceptual learning in visual hyperacuity. *ence*, 256(5059), 1018-1021.
- Shimono, K., Tam, W. J., & Nakamizo, S. (1999). Wheatstone—Panum limiting case: Occlusion, camouflage, and vergence-induced disparity cues. *Perception & Psychophysics*, 61(3), 445-455.
- Shiu, L. P., & Pashler, H. (1992). Improvement in line orientation discrimination is retinally local but dependent on cognitive set. *Percept Psychophys*, 52(5), 582-588.
- Somes, G. W., & Bhapkar, V. (1977). A simulation study on a Wald statistic and Cochran's Q statistic for stratified samples. *Biometrics*, 643-651.
- Spang, K., Gillam, B., & Fahle, M. (2012). Electrophysiological Correlates of Binocular Stereo Depth without Binocular Disparities. *Plos One*, 7.
- Tsirlin, I., Wilcox, L. M., & Allison, R. S. (2014). A computational theory of da Vinci stereopsis. *Journal of Vision*.
- Tsodyks, M., & Gilbert, C. (2004). Neural networks and perceptual learning. *Nature*, 431(7010), 775-781.
- Wang, Z., Wu, X. N., Ni, R., & Wang, Y. J. (2001). Double fusion does not occur in Panum's limiting case: evidence from orientation disparity. *Perception*, 30(9), 1143-1149.

- Wheatstone, C. (1838). Contributions to the Physiology of Vision.--Part the First. On Some Remarkable, and Hitherto Unobserved, Phenomena of Binocular Vision. *Philosophical Transactions of the Royal Society of London*, 128, 371-394.
- Wright, B. A., & Sabin, A. T. (2007). Perceptual learning: how much daily training is enough? *Experimental Brain Research*, 180(4), 727-736.
- Yu, Dongchuan, Jin, Laipeng, Li, Huayun, Jia, Huibin, Chung-Fat-Yim, & Ashley. (2017). The neural correlates of vertical disparity gradient and cue conflict in Panum's limiting case. *Brain & Cognition*.
- Zhou, Y., Huang, C., Xu, P., Tao, L., Qiu, Z., Li, X., & Lu, Z.-L. (2006). Perceptual learning improves contrast sensitivity and visual acuity in adults with anisometropic amblyopia. *Vision Research*, 46(5), 739-750.

## Figure Legends

Figure 1. An example of the monocular occlusions. The nearer surface differentially occludes the more distant surface in the two eyes, generating features that are visible to only one of the two eyes. In this figure, the red lines delimit the regions of the more distant surface visible to the left eye, and the blue lines delimit the regions visible to the right eye.

Figure 2. A schematic representation of Panum's limiting case from an arrangement of two real objects. The closer cube occludes the more distant cube for the left eye. But for the right eye, both cubes are visible. In this situation, the image presented to the left eye is only one cube since the two cubes are superimposed, while the image presented to the right eye are two cubes since the two cubes are located in different locations. The stimulus below each eye depicts the image that is presented in each eye's perspective.

Figure 3. An example of the stimulus used in this experiment. If a single straight line is presented to the left eye and two curved lines are presented to the right eye, observers would perceive the left 3D feature to appear closer than the right feature after binocular fusion. The other stimuli were varied by different vertical disparity gradient (0.1, 0.3 and 0.6) and cue conflict (low, medium and high) based on this stimulus.

Figure 4. Sequence of events within each trial. At the beginning of each trial, the reference plane consisted of two rectangles, which appeared for unlimited time. When the observers perceived the two rectangles fused into a single stable rectangle, they pressed the 'Space' key. The stimulus then appeared for 10 seconds. During this period, observers reported their perception of double fusion or one of the other percepts (single fusion, diplopia, or binocular rivalry).

Figure 5. The RTs results of vertical disparity gradient and cue conflict. The Y-axis represents the value of RTs to reported double fusion. Within each figure, from left to right, the three groups of columns represent the three different levels of vertical gradient (0.1, 0.3, and 0.6). Within each group, the three columns of colors represent low, medium and high cue conflict, respectively. Error bars represent standard errors of the means (SEM).

Figure 6. The RTs results for vertical disparity gradient and training. The X-axis represents the five training sessions with three levels of vertical gradient. The green, orange, and blue lines represent RTs of 0.1, 0.3 and 0.6 vertical disparity gradient, respectively. Error bars represent SEM.

Figure 7. The RTs results for cue conflict and training. The X-axis represents the five training sessions with three levels of conflict. The green, orange, and blue lines represent RTs of low, medium and high conflict, respectively. Error bars show SEM.

Figure 8. The duration results of double fusion for vertical disparity gradient and cue conflict. The Y-axis represents the duration to reported for double fusion. The other legends are the same as in Figure 5.

Figure 9. The duration results of double fusion for vertical gradient and training. The other legends are the same as in Figure 6.

Figures

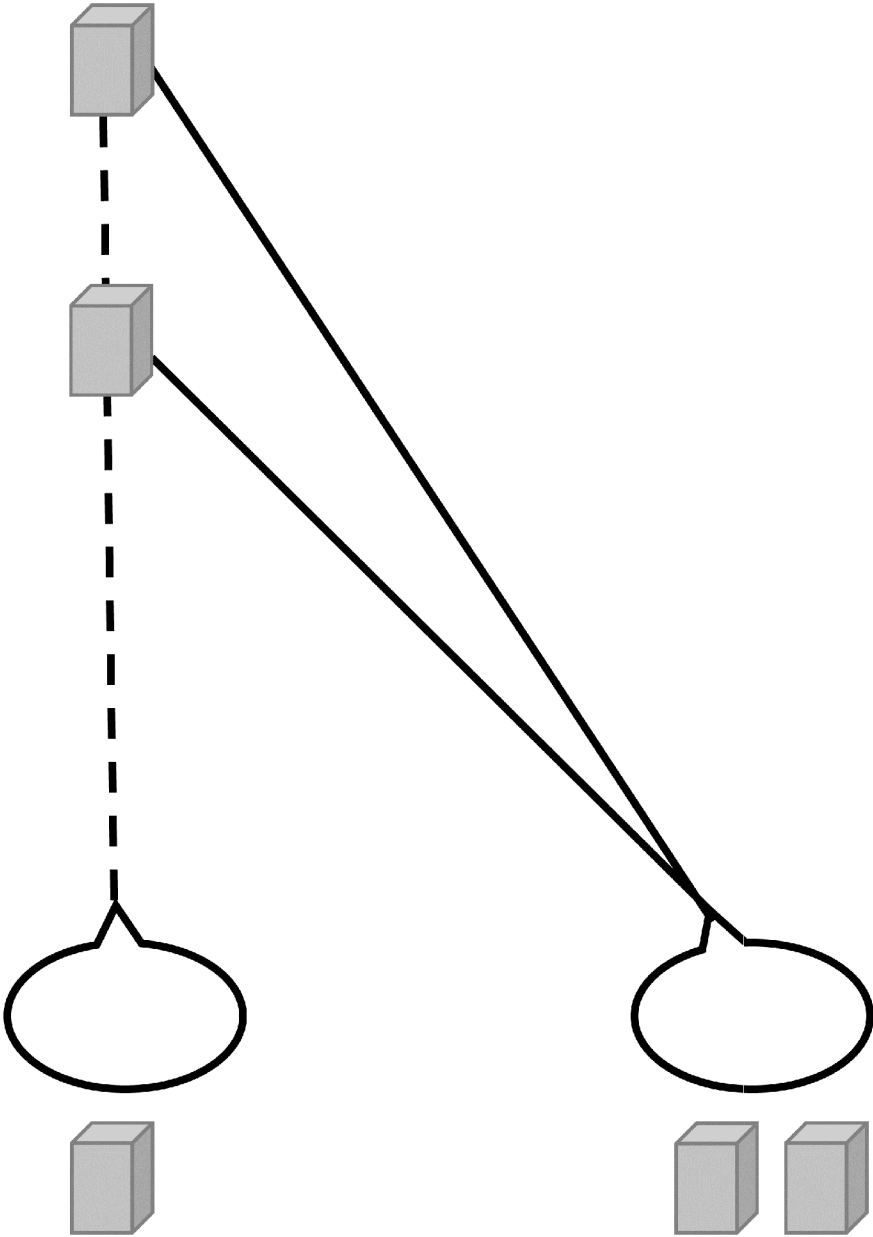


Figure 1.

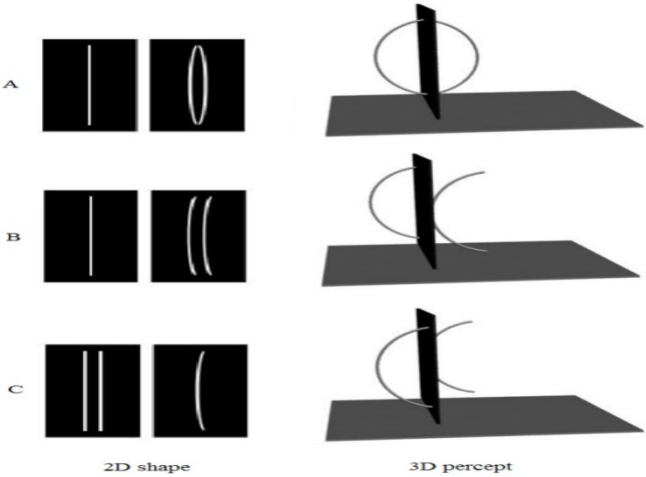


Figure 2.

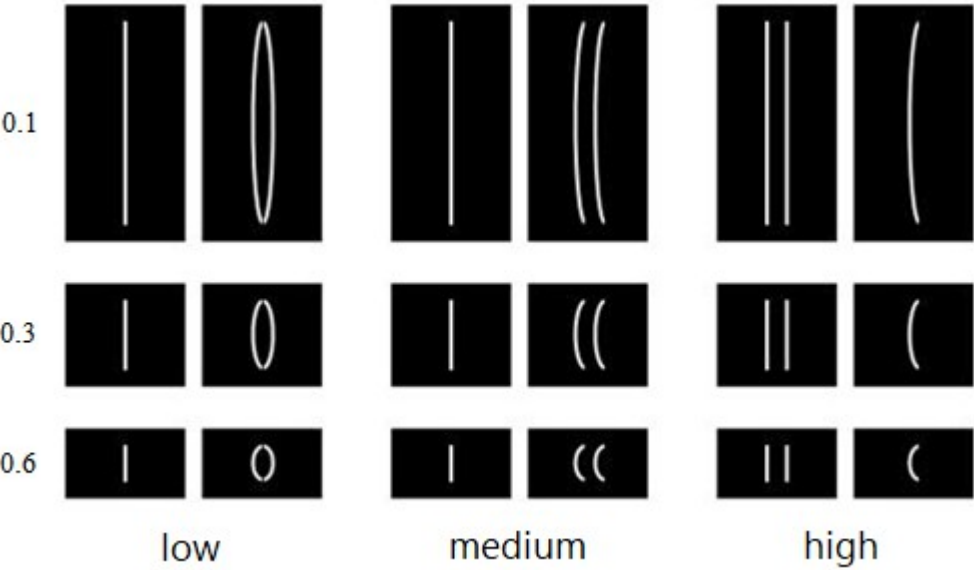


Figure 3.

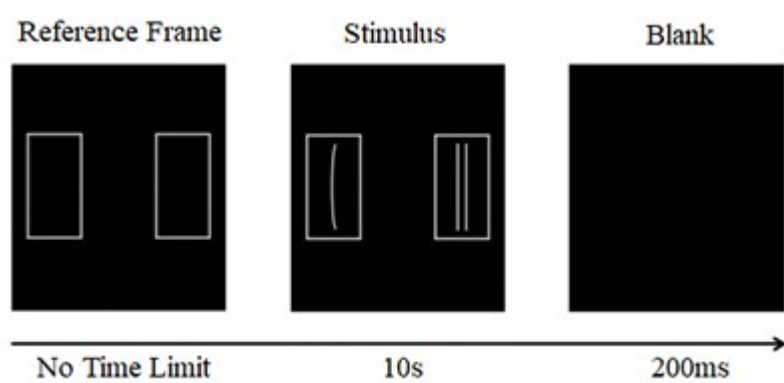


Figure 4.



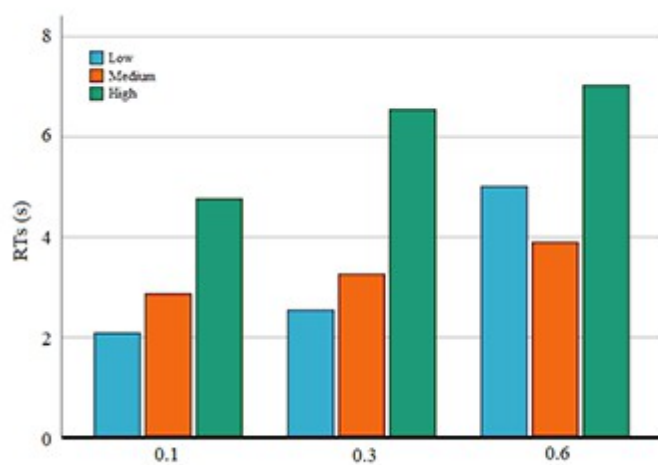


Figure 5.

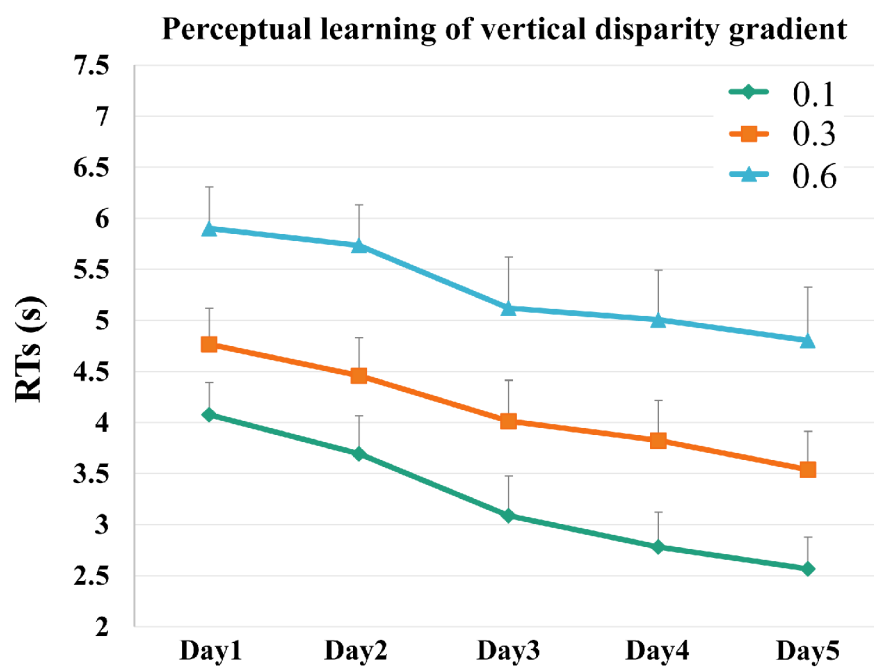


Figure 6.

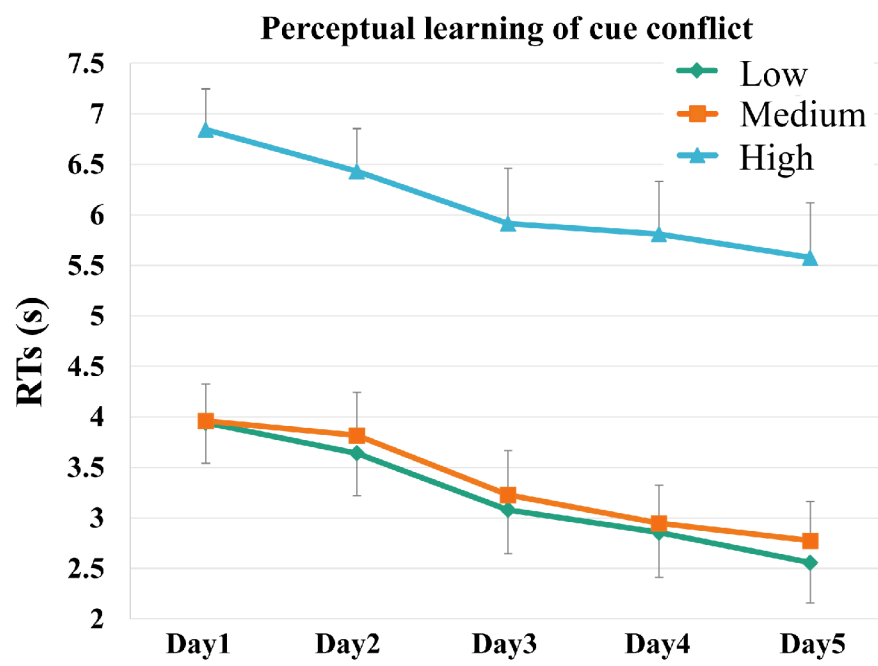


Figure 7.

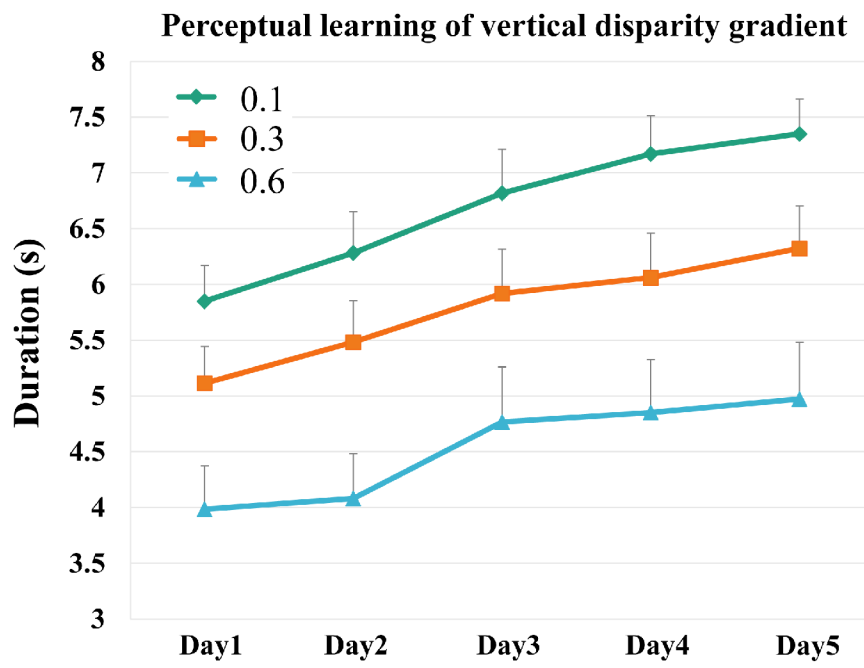


Figure 8.

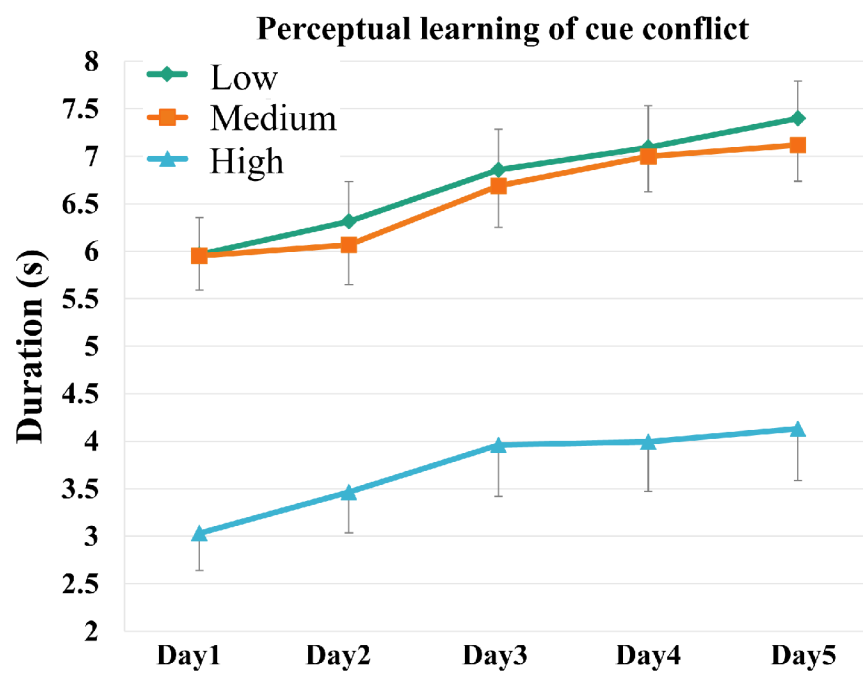


Figure 9.